Advances in Water Resources 82 (2015) 98–105

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

On the emergence of heavy-tailed streamflow distributions

S. Basso [a,b,*], M. Schirmer [a,b], G. Botter [c]

[a] Eawag – Swiss Federal Institute of Aquatic Science and Technology, Department of Water Resources and Drinking Water, Dübendorf, Switzerland
[b] Centre for Hydrogeology and Geothermics (CHYN), University of Neuchâtel, Neuchâtel, Switzerland
[c] Department ICEA and International Center for Hydrology "Dino Tonini", University of Padova, Padua, Italy

ARTICLE INFO

Article history:
Received 16 January 2015
Received in revised form 22 April 2015
Accepted 24 April 2015
Available online 29 April 2015

Keywords:
Streamflow distributions
High flows
Heavy-tails
Ecohydrology
Recession analysis
Storage–discharge

ABSTRACT

The right tail of streamflow distributions quantifies the occurrence probability of high flows, which play an important role in the dynamics of many eco-hydrological processes and eventually contribute to shape riverine environments. In this paper, the ability of a mechanistic analytical model for streamflow distributions to capture the statistical features of high flows has been investigated. The model couples a stochastic description of soil moisture dynamics with a simplified hydrologic response based on a catchment-scale storage–discharge relationship. Different types of relations between catchment water storage and discharge have been investigated, and alternative methods for parameter estimation have been compared using informal performance metrics and formal model selection criteria. The study highlights the pivotal role of non-linear storage–discharge relations in reproducing observed frequencies of high flows, and reveals the importance of analyzing the behavior of individual events for a reliable characterization of recession parameters. The emergence of heavy-tailed streamflow distributions is mechanistically linked to the degree of non-linearity of the catchment hydrologic response, with implications for the understanding of rivers' flooding potential and related ecologic and morphological processes.

1. Introduction

Heavy-tailed distributions are characteristic of many variables used in the description of natural and anthropogenic systems, including for example city populations and earthquake intensities [17]. These variables can assume values orders of magnitude greater than their averages, and are characterized by markedly skewed distributions which assign significant probabilities to extreme events [31].

Daily river flows have been previously recognized as a heavy-tail distributed variable [9,24]. Several statistical and physically-based models have been developed to characterize runoff dynamics and estimate streamflow distributions (e.g. [3,14,25,41]), and rainfall-runoff models have been specifically designed to incorporate heavy-tailed components (e.g. [12]).

Indeed, reliably modeling frequencies of high flows and identifying mechanisms promoting the emergence of heavy-tailed flow distributions entail important consequences for the characterization of a number of hydrological and ecological processes. For example, the enhanced frequencies of high flows associated with heavy-tailed distributions may result in wider areas of the river transect being affected by flooding, implying a wider aquatic/terrestrial transition zone within river corridors [40]. An accurate description of the tail of flow distributions is also important to characterize catchment-scale sediment transport, because of the role of high flows in mobilizing sediments and driving morphodynamic processes [39]. High flows are especially important for the assessment of formative discharges and in long-term landscape evolution models [38], and entail broad implications for different fields of geosciences that include reservoir sedimentation, geomorphology and river restoration [e.g. [35]]. Furthermore, the features of the tail of streamflow distributions may strongly impact the probabilistic structure of extreme flows, whose characterization is an important task for hydrologists and engineers. In fact, the extreme value theory postulates that different shapes of the tail of flow distributions result in different types of extremal distributions [20], which denote distinct characters of flooding. Hence, factors controlling the emergence of heavy tails should be identified to link expected changes of climate and land use to resulting modifications of flow frequency and magnitude.

This work, which represents a preliminary step towards a physically-based characterization of flood frequencies and catchment-scale sediment flow dynamics, aims at analyzing performances of a mechanistic model of streamflow dynamics [3,5]

* Corresponding author at: Eawag – Swiss Federal Institute of Aquatic Science and Technology, Department of Water Resources and Drinking Water, Überlandstrasse 133, P.O. Box 611, CH-8600, Dübendorf, Switzerland.

E-mail address: stefano.basso@eawag.ch (S. Basso).

http://dx.doi.org/10.1016/j.advwatres.2015.04.013

in the range of high flows. Recently, the model has been extensively applied to characterize river flow regimes and a variety of biotic and anthropogenic riverine processes (see e.g. [7,8,16,19,26,28,29,33,34]). The ability of the model to reproduce observed probability distributions in the range of low to medium streamflows has been the object of previous studies [4,15]. However, model performance for high flows has never been investigated before. In order to improve the model ability to reproduce the frequencies of high flows, a new method for parameter estimation is also tested. Moreover, conditions that promote the emergence of heavy tails in streamflow distributions are investigated. The paper is organized as follows: Section 2 describes the probabilistic characterization of streamflow dynamics, case studies and parameter estimation procedures. In Section 3, results of the application of the model to a set of catchments are described with a specific focus on the highest streamflows. The main conclusions of the paper are summarized in Section 4.

2. Methods

2.1. Probabilistic characterization of streamflows

The primary tool used in this investigation is the mechanistic-stochastic model for streamflow dynamics developed by [3,5]. The model builds on a catchment-scale balance of the soil moisture in the root zone, as resulting from the following processes: (1) stochastic increments due to infiltration from rainfall, which is assumed to be a Poisson process with frequency $\lambda_p[T^{-1}]$ and exponentially distributed depths with average $\alpha[L]$; (2) losses due to evapotranspiration; (3) instantaneous deep percolation above a certain soil moisture threshold. These deep percolation events, occurring with frequency $\lambda < \lambda_p[T^{-1}]$, recharge the subsurface storage of the catchment and produce effective rainfall which contributes to streamflow. The release of water is modeled by assuming a suitable storage–discharge relation which can be either linear or non-linear. In the former case, the dynamics of specific (per unit catchment area) streamflows q are described by:

$$\frac{dq}{dt} = -k\,q + \xi_1(t) \tag{1}$$

where $\xi_1(t)$ describes a sequence of random (exponentially distributed) streamflow increments due to percolation events and k is the hydrograph recession rate. In the latter case, instead, the dynamics of q are defined by:

$$\frac{dq}{dt} = -K\,q^a + \xi_2(t) \tag{2}$$

In analogy with Eq. (1), the term $\xi_2(t)$ in Eq. (2) represents (state dependent) streamflow increments due to percolation events; K and a (hereafter termed recession coefficient and exponent, respectively) represent the coefficient and exponent of the power law relation describing flow recessions.

Parameters k, K and a in Eqs. (1) and (2) not only describe recession patterns, but also influence the increase of streamflow following an effective rainfall event (i.e. ξ_1 and ξ_2). While in Eq. (1) a single catchment response time is assumed, thus implicitly referring to a specific component of the hydrologic response (e.g. subsurface runoff, see [3]), the non-linear model allows the catchment hydrologic response to vary as a function of the overall water storage. As a consequence, high saturation of the catchment results in higher streamflows, with no need to explicitly differentiate runoff generation mechanisms triggered by crossing of humidity (or rainfall) thresholds. Eq. (2) thus provides a consistent representation of the catchment hydrologic response, incorporating the effect of different flow components (deep, subsurface and surface runoff).

The model allows an analytical expression of the steady-state probability distribution (pdf) of specific streamflows, which reads [3,8]:

$$p(q) = \frac{\Gamma(\lambda/k)^{-1}}{\alpha k}\left(\frac{q}{\alpha k}\right)^{\frac{\lambda}{k}-1} \exp\left(-\frac{q}{\alpha k}\right) \tag{3}$$

in the linear case, and [5,15]:

$$p(q) = C\,q^{-a} \exp\left(-\frac{q^{2-a}}{\alpha K(2-a)} + \frac{\lambda\,q^{1-a}}{K(1-a)}\right) \tag{4}$$

in the non-linear case. C in Eq. (4) is a normalizing constant. The original formulation (see Eq. (2) of [15]) includes an atom of probability for $q = 0$, possibly arising when $0 < a < 1$, which has not been produced here because in most cases $a > 1$ [1].

The exceedance probability of q is obtained by integrating Eqs. (1) or (2) as:

$$D(q) = \int_q^{+\infty} p(x)dx \tag{5}$$

where $D(q)$ represents the flow duration curve at-a-station.

2.2. Case studies

The models described in Section 2.1 have been applied to estimate the seasonal distribution of streamflows and the related cumulative exceedance probability (i.e., the flow duration curve) in 16 catchments located in northeastern Italy, Switzerland and the United States, for a total of 43 different combinations of catchments and seasons. The selected basins are characterized by different flow regimes, and their areas span two orders of magnitude (from 10 to 10^3 km^2). To comply with the basic assumptions of the models, rivers affected by anthropogenic regulation as well as regimes significantly impacted by snow dynamics have been excluded from the analysis. Table 1 reports the main features of the investigated catchments.

2.3. Parameter estimation

The characterization of Eqs. (3) or (4) relies on the specification of three $(\alpha,\ \lambda,\ k)$ or four $(\alpha,\ \lambda,\ K,\ a)$ model parameters, which can be estimated based on hydrologic, climatic and geomorphologic information.

The mean rainfall depth α is estimated by using daily rainfall data recorded in several stations located within or nearby the considered basin. When synchronous data from different stations are available for the same catchment, rainfall records are first averaged to obtain spatially averaged rainfall series.

For the sake of simplicity, the frequency of effective rainfall λ is here estimated by equaling the observed mean specific discharge, $\langle q\rangle$, and the analytical mean of q according to the stochastic model (i.e., $\lambda = \langle q\rangle/\alpha$). An alternative estimation method exists, based solely on soil and climatic features of basins (see [8]).

Parameters defining the storage–discharge relations are set by using two alternative methods, whose outcomes are compared. (1) The ensemble of points derived by plotting time derivatives of q (dq/dt) against the corresponding flow values is fitted by a linear or power-law curve, from which the parameters k, a and K are estimated (dashed black line in Fig. 1). This is the standard procedure introduced by [10] and previously used by e.g. [4,5,8,15] to assess performance of the model described in Section 2.1. (2) A linear or power-law curve is fitted to pairs $dq/dt - q$ of each recession curve (colored lines in Fig. 1) to estimate k, a and K of individual recessions. By keeping a constant and equal to the median of the set of observed values, K is then re-calculated for each recession curve. The median values of the parameters across all the observed